

แนวคิดการวิเคราะห์กฎความสัมพันธ์ด้วย Apriori Algorithm เพื่อศึกษาปัญหาทางสังคมของเด็กและเยาวชน ด้อยโอกาสในประเทศไทย

อานนท์ สุกสัก และ พุดตาน พันธุ์เนตร

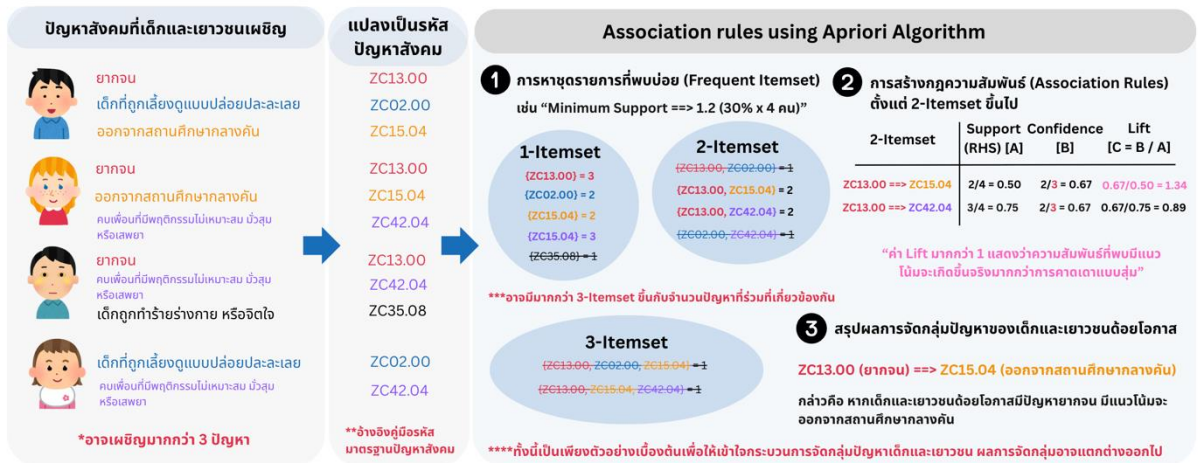
การวิเคราะห์กฎความสัมพันธ์ (Association Rule Mining) เป็นเทคนิคสำคัญในกระบวนการสืบค้นข้อมูล (Data Mining) ที่ได้รับการพัฒนาโดย Agrawal et al. (1993) ⁽¹⁾ มีเป้าหมายเพื่อค้นหารูปแบบหรือความสัมพันธ์ที่เกิดขึ้นร่วมกันอย่างมีนัยสำคัญภายในชุดข้อมูลขนาดใหญ่ กฎที่ได้จากกระบวนการนี้มักอยู่ในรูปแบบ “หากเกิดเหตุการณ์ A มักจะพบเหตุการณ์ B ร่วมด้วย” ซึ่งในทางธุรกิจมักใช้วิเคราะห์พฤติกรรมผู้บริโภค เช่น ผู้ซื้อสินค้าประเภทหนึ่งมักซื้อสินค้าประเภทอื่นควบคู่กัน ⁽²⁾ ทั้งนี้แนวคิดนี้สามารถนำมาประยุกต์ใช้กับข้อมูลด้านสังคม โดยเฉพาะพิจารณาความเกี่ยวข้องของปัญหาสังคมที่เกิดขึ้นกับเด็กและเยาวชนที่อยู่ในภาวะเปราะบาง จากการดำเนินการวิจัยที่ผ่านมาและการทบทวนวรรณกรรมพบว่า เด็กจำนวนมากมิได้เผชิญปัญหาเพียงด้านเดียว แต่มีความซับซ้อนที่เชื่อมโยงกัน เช่น เด็กที่มีฐานะยากจนมักมีแนวโน้มหลุดจากระบบการศึกษา และอาจประสบกับพฤติกรรมเสี่ยงหรือภาวะทางจิตสังคมร่วมด้วย การใช้เทคนิค Association Rule สามารถจำแนก “กลุ่มของปัญหาที่มักเกิดร่วมกัน” ทำให้มองเห็นภาพรวมของปัญหาได้ในลักษณะเครือข่าย มากกว่าการแยกวิเคราะห์เป็นรายประเด็น ซึ่งมีประโยชน์อย่างยิ่งในการวางแผนเชิงระบบ ออกแบบนโยบาย และพัฒนาระบบการเบิกจ่ายที่ทำให้เด็กและเยาวชนหลุดพ้นจากความด้อยโอกาสที่ยังเป็นความหวังของการแก้ปัญหาและความท้าทายในปัจจุบัน ⁽³⁾

1. ขั้นตอนการจัดกลุ่มปัญหาทางสังคมของเด็กและเยาวชนด้อยโอกาสโดยอาศัยการวิเคราะห์กฎความสัมพันธ์ด้วย Apriori Algorithm

การค้นหากฎความสัมพันธ์ในชุดข้อมูลที่มีรายการจำนวนมาก อาจต้องใช้ทรัพยากรมหาศาลหากดำเนินการแบบครบทุกความเป็นไปได้ ดังนั้นจึงมีการพัฒนาขั้นตอน Apriori Algorithm ขึ้น เพื่อช่วยลดภาระในการคำนวณ โดยใช้แนวคิด “หากชุดย่อยของข้อมูลไม่ปรากฏบ่อย ชุดข้อมูลที่มีขนาดใหญ่กว่าซึ่งประกอบด้วยชุดย่อยนั้นก็ย่อมไม่ปรากฏบ่อยเช่นกัน” กระบวนการทำงานของ Apriori สามารถสรุปเป็น 2 ขั้นตอนหลัก ⁽⁴⁾ ดังนี้

1. การหาชุดรายการที่พบบ่อย (Frequent Itemset) ขั้นตอนแรกเริ่มจากการค้นหารายการปัญหาเดียวที่พบได้บ่อยในกลุ่มเด็กและเยาวชน เช่น “ความยากจน” “ถูกปล่อยปละละเลย” หรือ “ออกจากสถานศึกษากลางคัน” จากนั้นนำรายการที่พบเหล่านี้มารวมกันเป็นชุดของปัญหาที่เกิดขึ้นร่วมกัน เช่น ชุดที่มี 2 รายการ (“ยากจน + ขาดเรียนบ่อย”) หรือ 3 รายการ และสามารถขยายจำนวนรายการในชุดได้เรื่อย ๆ ตามความถี่ของการเกิดร่วมกันในข้อมูล แต่ละชุดรายการจะต้องผ่านเกณฑ์ความถี่ขั้นต่ำที่เรียกว่า ค่าสนับสนุน (Support) เช่น หากพบว่ากลุ่มปัญหา “ถูกทำร้ายร่างกายหรือจิตใจ” หรือ “ยากจน + ถูกปล่อยปละละเลย” ปรากฏในเด็ก 1 คนจากทั้งหมด 4 คน จะมีค่า Support เท่ากับร้อยละ 25 หากค่า Support ที่พบต่ำกว่าระดับขั้นต่ำที่กำหนดไว้ (Minimum Support Threshold) เช่น น้อยกว่าร้อยละ 30 ระบบจะตัดชุดรายการนั้นทิ้งทันที หรือในกรณีที่พิจารณาเป็นจำนวนคน เช่น ฐานข้อมูลมีเด็ก 4 คน และกำหนด Minimum Support ไว้ที่ร้อยละ 30 จะเท่ากับ 1.2 คนหรือครึ่ง (30% x 4) ดังนั้นรายการใดที่มีความถี่เกิดซ้ำต่ำกว่า 1.2 คนหรือครึ่ง จะไม่ถูกนำมาพิจารณาในขั้นตอนถัดไป เฉพาะชุดรายการที่ผ่านเกณฑ์เท่านั้น จึงจะถูกนำไปใช้สร้างกฎความสัมพันธ์ โดยกฎจะต้องพิจารณาจากปัญหาที่เกิดขึ้นร่วมกัน ตั้งแต่ 2 รายการขึ้นไป ^(3, 5)

2. การสร้างกฎความสัมพันธ์ (Association Rules) จากชุดรายการที่ผ่านการคัดกรองในขั้นตอนแรก จะนำมาสร้างเป็นกฎในลักษณะ “หากพบ A แล้วมักพบ B” พร้อมคำนวณค่าตัวชี้วัด เช่น (1) ค่าความเชื่อมั่น (Confidence) ใช้ระบุความน่าจะเป็นที่ B จะเกิดขึ้นเมื่อ A เกิดขึ้น คำนวณจาก $Confidence(A \rightarrow B) = \frac{\text{จำนวนรายการที่มีทั้ง A และ B}}{\text{จำนวนรายการที่มี A}}$ และ (2) ค่าลิฟท์ (Lift) เป็นค่าที่สะท้อนความเข้มข้นของความสัมพันธ์ ว่าการพบ A แล้วพบ B เป็นความบังเอิญหรือมีแนวโน้มเกิดร่วมกันอย่างมีนัยสำคัญ คำนวณจาก $Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)}$ หากค่า Lift มากกว่า 1 แสดงว่าความสัมพันธ์ที่พบมีแนวโน้มจะเกิดขึ้นจริงมากกว่าการคาดเดาแบบสุ่ม ⁽⁵⁾ ซึ่งจะถูกนำไปพิจารณาเป็นกฎความสัมพันธ์ หรือเป็นกลุ่มปัญหาสังคมต่อไป ขั้นตอนนี้จะวนซ้ำโดยเพิ่มขนาดของ itemset ทีละระดับ พร้อมกรองทิ้งชุดที่ไม่ผ่านเกณฑ์ในทุกขั้นตอน จนกว่าจะไม่สามารถสร้างชุดใหม่ได้อีก กระบวนการนี้ช่วยให้ได้เฉพาะความสัมพันธ์ที่ “น่าเชื่อถือและมีนัยสำคัญ” เท่านั้น ^(2, 4, 5)



รูปที่ 1 ตัวอย่างการจัดกลุ่มปัญหาทางสังคมของเด็กและเยาวชนด้วยโอกาสโดยการวิเคราะห์กฎความสัมพันธ์ด้วย Apriori Algorithm

ที่มา: สังเคราะห์และดัดแปลงจากแนวคิดและข้อมูลของมูลนิธิศูนย์วิจัยและติดตามความเป็นธรรมทางสุขภาพ⁽³⁾ กัญญาวีร์ พรสว่างดี และณัฐพัชร์ เศรษฐศาสตร์⁽²⁾ และ ณัฐพัชร์ เศรษฐศาสตร์ และนัทธมน มยุระสาคร⁽⁵⁾ โดยผู้เขียน

2. การประยุกต์ใช้ Apriori Algorithm กับการจัดกลุ่มปัญหาทางสังคมของเด็กและเยาวชนด้วยโอกาสในประเทศไทย

การวิเคราะห์ Association Rule ร่วมกับ Apriori Algorithm มาใช้กับการจัดกลุ่มปัญหาเด็กและเยาวชน ช่วยให้สามารถจัด "กลุ่มปัญหาที่เกิดร่วมกันบ่อย" ได้อย่างเป็นระบบ เช่น กลุ่ม A เด็กที่เผชิญกับปัญหายากจน ขาดโอกาสทางการศึกษา และมีแนวโน้มพฤติกรรมเสี่ยง และ กลุ่ม B เด็กที่มีความรุนแรงในครอบครัวร่วมกับภาวะซึมเศร้าและไม่มีผู้ดูแลใกล้ชิด เป็นต้น การจัดกลุ่มเช่นนี้จะช่วยให้นักสังคมสงเคราะห์ หน่วยงานรัฐ หรือผู้กำหนดนโยบายสามารถออกแบบมาตรการช่วยเหลือที่ตรงตามลักษณะปัญหา เช่น การจัดสรรทุนการศึกษาให้กลุ่ม A หรือการส่งต่อกลุ่ม B ไปยังระบบสุขภาพจิตอย่างเหมาะสม⁽³⁾

ที่กล่าวมาเป็นแนวคิดเบื้องต้นในการจัดกลุ่มปัญหาสังคมสำหรับเด็กและเยาวชน ซึ่งเป็นก้าวสำคัญของการก้าวต่อการพิจารณาการเบิกจ่ายตามกลุ่มปัญหาสังคม ซึ่งขอเรียกการจัดกลุ่มนี้ว่า "กลุ่มปัญหาสังคมร่วม" อันเป็นข้อเสนอสำคัญของการพัฒนาระบบสร้างเสริมโอกาสให้เด็กและเยาวชนมีความเสมอภาคทางการศึกษาและสังคมต่อไป

เอกสารอ้างอิง

1. Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases. SIGMOD Rec. 1993;22(2):207-16.
2. กัญญาวีร์ พรสว่างดี, ณัฐพัชร์ เศรษฐศาสตร์. Association Rule: การหา Frequent Itemsets ด้วย Apriori Algorithm [อินเทอร์เน็ต]. กรุงเทพฯ: สถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน); 2564 [เข้าถึงเมื่อ 8 กรกฎาคม 2568]. เข้าถึงได้จาก: <https://bdi.or.th/big-data-101/frequent-itemsets-apriori-algorithm/>.
3. มูลนิธิศูนย์วิจัยและติดตามความเป็นธรรมทางสุขภาพ. รายงานฉบับสมบูรณ์โครงการพัฒนาระบบสร้างเสริมโอกาสให้เด็กมีความเสมอภาคทางการศึกษาและสังคม (ระยะที่ 1). กรุงเทพฯ: กองทุนเพื่อความเสมอภาคทางการศึกษา; 2567.
4. ไกรศักดิ์ เกษร. กฎความสัมพันธ์ (Association rule) [อินเทอร์เน็ต]. 2567 [เข้าถึงเมื่อ 9 กรกฎาคม 2568]. เข้าถึงได้จาก: <https://csit.u.ac.th/kraisak/ds/ds/chapter09/Chapter09.pdf>.
5. ณัฐพัชร์ เศรษฐศาสตร์, นัทธมน มยุระสาคร. มาทำความรู้จัก Association Rule: เครื่องมือเพื่อการวิเคราะห์ตะกร้าตลาด กัน! [อินเทอร์เน็ต]. กรุงเทพฯ: สถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน); 2564 [เข้าถึงเมื่อ 8 กรกฎาคม 2568]. เข้าถึงได้จาก: <https://bdi.or.th/big-data-101/what-is-association-rule/>.